First order strong approximations of scalar SDEs with values in a domain

Andreas Neuenkirch[*]   Lukasz Szpruch[†]

Abstract

We are interested in strong approximations of one-dimensional SDEs which have non-Lipschitz coefficients and which take values in a domain. Under a set of general assumptions we derive an implicit scheme that preserves the domain of the SDEs and is strongly convergent with rate one. Moreover, we show that this general result can be applied to many SDEs we encounter in mathematical finance and bio-mathematics. We will demonstrate flexibility of our approach by analysing classical examples of SDEs with sublinear coefficients (CIR, CEV models and Wright-Fisher diffusion) and also with superlinear coefficients (3/2-volatility, Ait-Sahalia model). Our goal is to justify an efficient Multi-Level Monte Carlo (MLMC) method for a rich family of SDEs, which relies on good strong convergence properties.

Key words: *Stochastic differential equations, non-Lipschitz coefficients, Lamperti Transformation, backward Euler-Maruyama scheme.*

2000 Mathematics Subject Classification: 60H10, 65J15

1   Introduction

The goal of this paper is to derive an efficient numerical approximation for one-dimensional SDEs which take values in a domain and have non-Lipschitz drift or diffusion coefficients. Typical examples of such SDEs are the Cox-Ingersoll-Ross process (CIR), the CEV model, the Wright-Fisher diffusion, where the main difficulty is the sublinearity of the diffusion coefficient. Furthermore, the approach developed in this paper can be also applied to SDEs with superlinear coefficients. Prominent examples are here the Heston 3/2-volatility process and the Ait-Shalia model. All the mentioned processes play an important role in mathematical finance and bio-mathematical applications. Our key idea is to transform the original SDE using the Lamperti transformation into a SDE with constant diffusion coefficient, see e.g. [18]. The transformed SDE is then approximated by a backward (also called drift-implicit) Euler-Maruyama scheme (BEM) and transforming back yields an approximation scheme for the original SDE. This strategy was found successful in a recent work [6] for the CIR process, where the authors proved that the piecewise linearly interpolated BEM scheme strongly converges with rate one half (up to a log-term) with respect to a uniform

[*]Institut für Mathematik, Universität Mannheim, A5, 6, D-68131 Mannheim, Germany, neuenkirch@kiwi.math.uni-mannheim.de

[†]Mathematical Institute, University of Oxford, 24-29 St Giles', Oxford OX1 3LB, England Szpruch@maths.ox.ac.uk

L^p-error criterion. This strategy was also suggested by Alfonsi in [1]. Here, we extend that work in several ways:

- Considering the maximum error in the discretization points, we prove that the drift-implicit Euler-Maruyama scheme for the CIR process strongly converges with rate one under slightly more restrictive conditions on the parameters of the process than in [6].

- We provide a general framework for the strong order one convergence of the BEM scheme for SDEs with constant diffusion and one-sided Lipschitz drift coefficients.

- Using this framework we present a detailed convergence analysis for several SDEs with sub- and super-linear coefficients.

- We also show that BEM for the transformed SDE is closely related to a drift-implicit Milstein scheme for the original SDE, which has been introduced in [13]. In the case of the CIR process we provide a sharp error analysis for this scheme.

Independently of and simultaneously to the research presented in this paper, the same approach was also used by Alfonsi in [2] to derive strong order one convergence of the BEM scheme for the CIR and the CEV process. See Remark 2.9 for a discussion.

To illustrate the main difficulties and also our main idea let us consider the CIR process

$$dy(t) = \kappa(\theta - y(t))dt + \sigma\sqrt{y(t)}dw(t) \tag{1}$$

with $2\kappa\theta \geq \sigma^2$. It is a simple implication of the Feller test that the solution of equation (1) is strictly positive when $2\kappa\theta \geq \sigma^2$ and $y(0) > 0$. This SDE is often used in mathematical finance for interest rate or stochastic volatility models. However scalar SDEs with square root diffusion coefficients appear not only in the financial literature but belong to the most fundamental SDEs as they are an approximation to Markov jump processes [7].

Once we attempt to simulate (1) using classical discretization methods, see e.g. [24], we face two difficulties:

- In general, these methods do not preserve positivity and therefore are not well defined when directly applied to equation (1);

- The diffusion term is not Lipschitz continuous and therefore standard assumptions required for weak and strong convergence, see e.g. [24], do not hold.

Consequently, a considerable amount of research was devoted to the numerical approximation of this equation, see [11, 1, 21, 4, 26, 3], to mention a few. However no strong convergence of order one results have been obtained so far up to best of our knowledge. For a comparison of the different proposed schemes based on simulation studies, see [1, 26].

Our approach is based on a suitable transformation of the CIR process. Applying Itô's formula to $x(t) = \sqrt{y(t)}$ gives

$$dx(t) = \frac{1}{2}\kappa\left(\left(\theta - \frac{\sigma^2}{4\kappa}\right)x(t)^{-1} - x(t)\right)dt + \frac{1}{2}\sigma dw(t). \tag{2}$$

Zhu [31] pointed out that the drawback of the transformed equation is that the new mean level $(\theta - \frac{\sigma^2}{4\kappa})x(t)^{-1}$ is stochastic and that a naive Euler discretization cannot capture the erratic behavior of $x(t)^{-1}$-term, although almost sure convergence of this method holds true, see [19]. The weakness of a naive Euler discretization is that its transition density is Gaussian and therefore its moments explode due to the $x(t)^{-1}$-term.

On the other hand, Alfonsi showed in [1] that the BEM applied to (1) preserves positivity of the solution and also monotonicity with respect to the initial value. Moreover, his simulation studies indicated good convergence properties of this scheme. In this paper we follow the recent result by Dereich et al. [6], where it was shown that the piecewise linear interpolation of BEM applied to (1), strongly converges with a rate one half (up to a log-term).

Given any step size $\Delta t > 0$, the BEM scheme has the form

$$
\begin{aligned}
X_{k+1} &= X_k + \frac{1}{2}\kappa\big(\theta_v X_{k+1}^{-1} - X_{k+1}\big)\Delta t + \frac{1}{2}\sigma\Delta w_{k+1}, \qquad k = 0, 1, \ldots, \quad (3) \\
X_0 &= x(0)
\end{aligned}
$$

with

$$
\Delta w_{k+1} = w((k+1)\Delta t) - w(k\Delta t), \qquad k = 0, 1, \ldots
$$

and

$$
\theta_v = \theta - \frac{\sigma^2}{4\kappa}.
$$

We will establish a strong convergence of order one for the maximum L^p-distance in the discretization points between (1) and (3), see Section 3. For example for the L^2-distance we obtain

$$
\mathbb{E}\max_{k=0,\ldots,\lceil T/\Delta t\rceil}|X_k - x(k\Delta t)|^2 \leq C_2 \cdot \Delta t^2 \qquad \text{for} \quad \frac{2\kappa\theta}{\sigma^2} > 3.
$$

As a consequence we also obtain the same convergence order for the approximation of the original CIR process by X_k^2, $k = 0, 1, \ldots, \lceil T/\Delta t\rceil$.

In this paper we will show that the above idea naturally extends to many types of SDEs with non-Lipschitz coefficients. The combination of the Lamperti transformation and the backward Euler scheme enables us to analyse the L^p-convergence rates for many scalar SDEs encountered in practice. In particular, transforming BEM back we obtain an order one scheme for the original SDE, which is close to a Milstein-type scheme, see Section 4. Hence our approach turns out to be a new method for deriving numerical methods with strong order one convergence for SDEs with non-Lipschitz coefficients. Although strong convergence of backward schemes for SDEs with non-Lipschitz coefficients has already been analysed in the literature and their convergence for models as the Ait-Sahalia and the Heston 3/2-volatility was obtained, see [13, 27, 29], schemes with strong convergence order one have not been established yet in this setting.

Another motivation for our work are results by Giles [8, 9], who showed that for optimal MLMC simulations one should use discretization schemes with strong convergence order one. Note that strong convergence of the discretization scheme used for the MLMC simulations seems to be not only a sufficient but also a necessary condition [16].

The remainder of this paper is structured as follows. In the next section, we provide a general convergence result for the BEM method applied to scalar SDEs with additive noise. Section 3 contains the results for our examples, i.e. the CIR, CEV, Ait-Sahalia, 3/2-Heston volatility and Wright-Fischer SDEs. In Section 4, we provide the relation of BEM and a drift-implicit Milstein scheme and give an error analysis for the case of the CIR process. The last section contains a short discussion.

2 The BEM scheme for SDEs with additive noise

2.1 Preliminaries

Let $D = (l, r)$, where $-\infty < l < r < \infty$, and let $a, b : D \to D$ be continuously differentiable functions. Moreover, let $(\Omega, \mathcal{F}, (\mathcal{F}_t)_{t \geq 0}, \mathbb{P})$ be a filtered probability space and $w(t)$, $t \geq 0$, a standard $(\mathcal{F}_t)_{t \geq 0}$-Brownian motion. We begin with the SDE

$$dy(t) = a(y(t))dt + b(y(t))dw(t), \quad t \geq 0, \qquad y(0) \in D, \qquad (4)$$

and assume that it has a unique strong solution with

$$\mathbb{P}(y(t) \in D, \ t \geq 0) = 1.$$

If $b(x) > 0$ for all $x \in D$, then we can use the Lamperti-type transformation

$$F(x) = \lambda \int^x \frac{1}{b(y)} dy$$

for some $\lambda > 0$. Itô's Lemma with $x(t) = F(y(t))$ gives the transformed SDE

$$dx(t) = f(x(t))dt + \lambda dw(t), \quad t \geq 0, \qquad x(0) \in F(D),$$

with

$$f(x) = \lambda \left(\frac{a(F^{-1}(x))}{b(F^{-1}(x))} - \frac{1}{2}b'(F^{-1}(x)) \right), \qquad x \in F(D),$$

where $F(D) = (F(l), F(r))$. Note that the classical Lamperti transformation corresponds to $\lambda = 1$. This transformation allows to shift non-linearities from the diffusion coefficient into the drift coefficient. Then, under appropriate assumptions on f (respectively a and b), one can apply the backward Euler scheme

$$X_{k+1} = X_k + f(X_{k+1})\Delta t + \lambda \Delta w_{k+1} \qquad (5)$$

and derive a strong convergence rate of order one for the maximum L^p-error in the discretization points, see Theorem 2.7 in the in Section 2.3.

2.2 Backward Euler-Maruyama scheme

In this section we focus on the numerical approximation of

$$dx(t) = f(x(t))dt + \sigma \, dw(t), \quad t \geq 0, \qquad x(0) = x_0 \tag{6}$$

by the backward Euler-Maruyama scheme

$$X_{k+1} = X_k + f(X_{k+1})\Delta t + \sigma\Delta w_{k+1}, \quad k = 0, 1, \ldots, \qquad X_0 = x_0. \tag{7}$$

We will work under the following assumption on the SDE itself:

Assumption 2.1. *Let* $-\infty < \alpha < \beta < \infty$ *and assume that SDE* (6) *has a unique strong solution which takes values in the set* $(\alpha, \beta) \subset \mathbb{R}$, *i.e.*

$$\mathbb{P}(x(t) \in (\alpha, \beta), \ t \geq 0) = 1. \tag{8}$$

For the well-definedness of the drift-implicit Euler-Maruyama scheme we need the following assumption on the drift-coefficient:

Assumption 2.2. *The function* $f : (\alpha, \beta) \to (\alpha, \beta)$ *is continuous. Moreover, there exists a constant* $K \in \mathbb{R}$ *such that*

$$(x - y)(f(x) - f(y)) \leq K \, |x - y|^2 \tag{9}$$

for all $x, y \in (\alpha, \beta)$.

The Feller test, see e.g. Theorem V.5.29 in [22], gives that condition (8) is equivalent to

$$\lim_{x \to \alpha_+} v(x) = \infty, \qquad \lim_{x \to \beta_-} v(x) = \infty$$

with

$$v(x) = \frac{2}{\sigma^2} \int_{x_0}^x \int_{x_0}^{\zeta_2} \frac{p'(\zeta_2)}{p'(\zeta_1)} d\zeta_1 d\zeta_2$$

and the scale function

$$p(x) = \int_{x_0}^x \exp\left(-2 \int_{x_0}^\xi \frac{f(u)}{\sigma^2} \, du\right) d\xi, \qquad x \in (\alpha, \beta).$$

Note that v can be rewritten as

$$v(x) = \frac{2}{\sigma^2} \int_{x_0}^x \int_{x_0}^{\zeta_2} \exp\left(-2 \int_{\zeta_1}^{\zeta_2} \frac{f(u)}{\sigma^2} \, du\right) d\zeta_1 d\zeta_2.$$

The condition on v now implies (recall that $\alpha < x_0 < \beta$)

$$\lim_{x \to \alpha_+} \int_x^{x_0} \exp\left(2 \int_x^{\zeta_1} \frac{f(u)}{\sigma^2} \, du\right) d\zeta_1 = \infty \tag{10}$$

and

$$\lim_{x \to \beta_-} \int_{x_0}^x \exp\left(-2 \int_{\zeta_1}^x \frac{f(u)}{\sigma^2} \, du\right) d\zeta_1 = \infty. \tag{11}$$

Now consider (10) and assume that $\limsup_{x \to \alpha_+} f(x) \neq \infty$ However, if this would be true, the expression in (10) would be finite due to the continuity of f. Using a similar argument for (11) we obtain

$$\limsup_{x \to \alpha_+} f(x) = \infty, \qquad \liminf_{x \to \beta_-} f(x) = -\infty. \qquad (12)$$

The drift-implicit Euler scheme is well defined if

$$X_{k+1} - f(X_{k+1})\Delta t = X_k + \sigma \Delta w_{k+1}$$

has a unique solution for $k = 0, 1, \ldots$. This is guaranteed by the following result:

Lemma 2.3. *Let Assumption 2.1 and 2.2 hold and let $K\Delta t < 1$. Moreover set*

$$G(x) = x - f(x)\Delta t, \quad x \in (\alpha, \beta).$$

Then for any $c \in \mathbb{R}$ there exists a unique $x \in (\alpha, \beta)$ such that $G(x) = c$.

Proof. The result follows if we can show that the function G is continuous, coercive and strictly monotone on (α, β) (see [30]). However, due to Assumption 2.2 the function G is continuous on (α, β). Moreover, since

$$(x - y)(G(x) - G(y)) = (x - y)^2 - (x - y)(f(x) - f(y))\Delta t$$
$$\geq (1 - K^+ \Delta t)(x - y)^2 > 0$$

by (9) (with $K^+ = \max\{0, K\}$) the required strict monotonicity is obtained. Finally, (12) and the monotonicity imply that

$$\liminf_{x \to \alpha_+} G(x) = \lim_{x \to \alpha_+} G(x) = -\infty$$

and

$$\limsup_{x \to \beta_-} G(x) = \lim_{x \to \beta_-} G(x) = \infty,$$

so the function G is coercive on (α, β). $\qquad \square$

Note that for $K \leq 0$ there is no restriction on Δt.

For completeness, we state here a well known discrete version of Gronwall's Lemma:

Lemma 2.4. *Let $\Delta t > 0$ and let $g_n, \lambda_n \in \mathbb{R}$, $n \in \mathbb{N}$, and $\eta \geq 0$ be given. Moreover, assume that $1 - \eta\Delta t > 0$ and $1 + \lambda_n > 0$, $n \in \mathbb{N}$. Then, if $a_n \in \mathbb{R}$, $n \in \mathbb{N}$, satisfies $a_0 = 0$ and*

$$a_{n+1} \leq a_n(1 + \lambda_n) + \eta a_{n+1}\Delta t + g_{n+1}, \qquad n = 0, 1, \ldots,$$

then this sequence also satisfies

$$a_n \leq \frac{1}{(1 - \eta\Delta t)^n} \sum_{j=0}^{n-1} (1 - \eta\Delta t)^j g_{j+1} \prod_{l=j+1}^{n-1} (1 + \lambda_l), \qquad n = 0, 1, \ldots.$$

Under the above assumptions we have the following moment bounds for the SDE and the BEM scheme:

Lemma 2.5. *Let $T > 0$ and let Assumption 2.1 and 2.2 hold. Then we have*

$$\mathbb{E} \sup_{t \in [0,T]} |x(t)|^q < \infty$$

for all $q \geq 1$. If additionally $2K\Delta t < \eta$ for some $\eta < 1$, then for all $q \geq 1$ there exist constants $C_q > 0$, which are independent of Δt, such that

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |X_k|^q \leq C_q.$$

Proof. The first assertion can be shown by a straightforward modification of the proof of Lemma 3.2 in [12].

For the proof of the second assertion, we will denote constants which are independent of Δt and whose particular value is not important by c regardless of their value. Due to (12) and the continuity of f there exists an $x^* \in (\alpha, \beta)$ with $f(x^*) = 0$, so we can rewrite BEM as

$$X_{k+1} - x^* = X_k - x^* + (f(X_{k+1}) - f(x^*))\Delta t + \sigma \Delta w_{k+1}.$$

Multiplying with $X_{k+1} - x^*$ and using the one-sided Lipschitz condition on f yield that

$$(X_{k+1} - x^*)^2 \leq ((X_k - x^*) + \sigma \Delta w_{k+1})(X_{k+1} - x^*) + K^+ \Delta t (X_{k+1} - x^*)^2.$$

where $K^+ = \max\{0, K\}$. Moreover, using $ab \leq \frac{1}{2}(a^2 + b^2)$ and rearranging the terms gives

$$(X_{k+1} - x^*)^2 \leq 2K^+ \Delta t (X_{k+1} - x^*)^2 + ((X_k - x^*) + \sigma \Delta w_{k+1})^2$$

and hence, with $u_k = X_k - x^*$, we have

$$u_{k+1}^2 \leq u_k^2 + 2K^+ u_{k+1}^2 \Delta t + 2u_k \Delta w_{k+1} + \sigma^2 |\Delta w_{k+1}|^2.$$

Note that $1 - 2K^+ \Delta t \in (0, 1]$ and

$$\sup_{\Delta t \in (0, \eta/(2K^+))} \sup_{k=1,\ldots,\lceil T/\Delta t\rceil} (1 - 2K^+ \Delta t)^{-k} < \infty. \tag{13}$$

So, the above discrete version of Gronwall's Lemma now yields

$$u_k^2 \leq c + \frac{2}{(1 - 2K^+ \Delta t)^k} \underbrace{\sum_{j=0}^{k-1} (1 - 2K^+ \Delta t)^j u_j \Delta w_{j+1}}_{=M_k^{(1)}} \tag{14}$$

$$+ \frac{1}{(1 - 2K^+ \Delta t)^k} \underbrace{\sum_{j=0}^{k-1} (1 - 2K^+ \Delta t)^j \sigma^2 \left(|\Delta w_{j+1}|^2 - \Delta\right)}_{=M_k^{(2)}},$$

from which we obtain easily by induction that

$$\sup_{k=0,\ldots,\lceil T/\Delta\rceil} \mathbb{E} u_k^2 < \infty.$$

Using this it can be easily checked that the processes $M_k^{(i)}$, $i = 1, 2$, are square-integrable martingales with respect to the filtration $\mathcal{F}_{k\Delta t}$, $k = 0, 1, \ldots$. Hence Doob's inequality and straightforward calculations give for any $q \geq 1$ that

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta \rceil} |M_k^{(1)}|^{2q} \leq c \cdot \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} \mathbb{E}|u_k|^{2q}$$

and

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |M_k^{(2)}|^{2q} < \infty.$$

So using (14) and (13) we obtain

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} u_k^{4q} \leq c + c \cdot \sup_{k=0,\ldots,\lceil T/\Delta t \rceil} \mathbb{E}u_k^{2q}$$

for $\Delta t < \eta/(2K^+)$ and the assertion follows now by an induction argument in $q \in \mathbb{N}$. $\square$

2.3 The Main Result

Here we prove our general theorem on the strong convergence of the numerical scheme (7) to the solution of SDE (6).

Assumption 2.6. *Let $T > 0$ and $p \geq 2$. We assume that the drift coefficient $f : (\alpha, \beta) \to (\alpha, \beta)$ of SDE (6) is twice continuously differentiable and satisfies*

$$\sup_{t \in [0,T]} \mathbb{E}\,|f'(x(t))|^p + \sup_{t \in [0,T]} \mathbb{E} \left| (f'f)(x(t)) + \frac{\sigma^2}{2} f''(x(t)) \right|^p < \infty.$$

Theorem 2.7. *Let $T > 0$, $p \geq 2$, $\eta \in (0,1)$ and Assumptions 2.1, 2.2 and 2.6 hold. Then, for $2K\Delta t < \eta$, there exists a constant $C_p > 0$ (independent of Δt) such that*

$$\mathbb{E} \left[\sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |x(k\Delta t) - X_k|^p \right] \leq C_p \cdot \Delta t^p. \tag{15}$$

Proof. Recall that we will denote constants which are independent of Δt and whose particular value is not important by c regardless of their value. By applying Itô's formula on $f(x(t))$ we have

$$x((k+1)\Delta t) = x(k\Delta t) + \int_{k\Delta t}^{(k+1)\Delta t} f(x((k+1)\Delta t))dt \tag{16}$$

$$+ \sigma \int_{k\Delta t}^{(k+1)\Delta t} dw(t) + R_{k+1}$$

where

$$R_{k+1} = -\int_{k\Delta t}^{(k+1)\Delta t} \int_t^{(k+1)\Delta t} \left((f'f)(x(s)) + \frac{\sigma^2}{2} f''(x(s)) \right) ds\,dt \tag{17}$$

$$- \sigma \int_{k\Delta t}^{(k+1)\Delta t} \int_t^{(k+1)\Delta t} f'(x(s))dw(s)dt.$$

8

We can decompose R_s as $R_s = R_s^{(1)} + R_s^{(2)}$ with

$$R_s^{(1)} = -\int_{(s-1)\Delta t}^{s\Delta t} \int_t^{s\Delta t} \left((f'f)(x(u)) + \frac{\sigma^2}{2} f''(x(u)) \right) du\, dt$$

$$R_s^{(2)} = -\sigma \int_{(s-1)\Delta t}^{s\Delta t} \int_t^{s\Delta t} f'(x(u)) dw(u) dt.$$

Using equations (16) and (7) we have

$$x((k+1)\Delta t) - X_{k+1} = x(k\Delta t) - X_k + [f(x((k+1)\Delta t)) - f(X_{k+1})]\Delta t + R_{k+1}$$

and thus

$$\big(x((k+1)\Delta t) - X_{k+1} - [f(x((k+1)\Delta t)) - f(X_{k+1})]\Delta t\big)^2 = \big(x(k\Delta t) - X_k + R_{k+1}\big)^2.$$

We arrive at

$$\begin{aligned}
|x((k+1)\Delta t) &- X_{k+1}|^2 - |x(k\Delta t) - X_k|^2 \\
&= 2(x((k+1)\Delta t) - X_{k+1})[f(x((k+1)\Delta t)) - f(X_{k+1})]\Delta t \\
&\quad - [f(x((k+1)\Delta t)) - f(X_{k+1})]^2 \Delta t^2 + 2(x(k\Delta t) - X_k)R_{k+1} + R_{k+1}^2.
\end{aligned}$$

Using the one-sided Lipschitz assumption on f we obtain

$$(1 - 2K^+ \Delta t)\, |x((k+1)\Delta t) - X_{k+1}|^2 \leq |x(k\Delta t) - X_k|^2 + 2(x(k\Delta t) - X_k)R_{k+1} + R_{k+1}^2.$$

Let us define $e_k = x(k\Delta t) - X_k$ and $\gamma_{\Delta t} = (1 - 2K^+ \Delta t)$. Note that $\gamma_{\Delta t} \in (0, 1]$ and

$$\sup_{\Delta t \in (0, \eta/(2K^+))} \sup_{k=1,\ldots,\lceil T/\Delta t \rceil} \gamma_{\Delta t}^{-k} < \infty. \tag{18}$$

Now, Lemma 2.4 yields

$$e_k^2 \leq 2 \sum_{s=0}^{k-1} \gamma_{\Delta t}^{s-k} e_s R_{s+1} + \sum_{s=0}^{k-1} \gamma_{\Delta t}^{s-k} R_{s+1}^2. \tag{19}$$

Since $\mathbb{E}\big[R_{s+1}^{(2)} \big| \mathcal{F}_{s\Delta t}\big] = 0$, we have that

$$\sum_{s=0}^{k-1} \gamma_{\Delta t}^s e_s R_{s+1}^{(2)}, \qquad k = 0, \ldots, \lceil T/\Delta t \rceil,$$

is a martingale and the Burkholder-Davis-Gundy inequality implies that

$$\mathbb{E}\left[\sup_{k=0,\ldots,\ell} \left| \sum_{s=0}^{k-1} \gamma_{\Delta t}^s e_s R_{s+1}^{(2)} \right|^q \right] \leq c\, \mathbb{E} \left(\sum_{s=0}^{\ell-1} \gamma_{\Delta t}^{2s} |e_s|^2 \left| R_{s+1}^{(2)} \right|^2 \right)^{q/2}$$

for any $q \geq 1$ and $\ell = 0, \ldots, \lceil T/\Delta t \rceil$. So using the boundedness of $\gamma_{\Delta t}$ and (18) we have

$$\mathbb{E}\left[\sup_{k=0,\ldots,\ell} \left| \sum_{s=0}^{k-1} \gamma_{\Delta t}^{s-k} e_s R_{s+1}^{(2)} \right|^q \right] \leq c\, \mathbb{E} \left(\sum_{s=0}^{\ell-1} |e_s|^2 \left| R_{s+1}^{(2)} \right|^2 \right)^{q/2}$$

9

for any $q \geq 1$. Using this and Jensen's inequality in (19) we now arrive at

$$
\mathbb{E}\left[\sup_{k=0,\ldots,\ell} |e_k|^{2q}\right] \leq c \lceil T/\Delta t\rceil^{q/2-1} \sum_{s=0}^{\ell-1} \mathbb{E}|e_s|^q \left|R_{s+1}^{(2)}\right|^q \tag{20}
$$

$$
+ c \lceil T/\Delta t\rceil^{q-1} \sum_{s=0}^{\ell-1} \mathbb{E}|e_s|^q \left|R_{s+1}^{(1)}\right|^q
$$

$$
+ c \lceil T/\Delta t\rceil^{q-1} \sum_{s=0}^{\ell-1} \mathbb{E}|R_{s+1}|^{2q}.
$$

Now, Assumption 2.6, Jensen's inequality and the Burkholder-Davis-Gundy inequality give that

$$
\mathbb{E}\left[\left|R_{k+1}^{(2)}\right|^m \Big| \mathcal{F}_{k\Delta t}\right] \leq c \left(1 + |f'(x_{k\Delta})|^m\right) \Delta t^{3m/2} \tag{21}
$$

and

$$
\mathbb{E}\left[\left|R_{k+1}^{(1)}\right|^m \Big| \mathcal{F}_{k\Delta t}\right] \leq c \left(1 + \left|(f'f)(x(k\Delta)) + \frac{\sigma^2}{2} f''(x(k\Delta))\right|^m\right) \Delta t^{2m} \tag{22}
$$

for all $m \leq p$. Thus, the Cauchy-Schwarz inequality and Assumption 2.6 yield that

$$
\mathbb{E}|e_s|^q \left|R_{s+1}^{(2)}\right|^q = \mathbb{E}\left[|e_s|^q \, \mathbb{E}\left[\left|R_{s+1}^{(2)}\right|^q \Big| \mathcal{F}_{s\Delta t}\right]\right] \leq c \left(1 + \mathbb{E}|e_s|^{2q}\right)^{1/2} \Delta t^{3q/2}.
$$

Hence Young's inequality implies

$$
c \lceil T/\Delta t\rceil^{q/2-1} \sum_{s=0}^{\ell-1} \mathbb{E}|e_s|^q \left|R_{s+1}^{(2)}\right|^q \leq c \sum_{s=0}^{\ell-1} \mathbb{E}\,|e_s|^{2q}\, \Delta t + c\Delta t^{2q}.
$$

Similar we also obtain

$$
c \lceil T/\Delta t\rceil^{q-1} \sum_{s=0}^{\ell-1} \mathbb{E}|e_s|^q \left|R_{s+1}^{(1)}\right|^q \leq c \sum_{s=0}^{\ell-1} \mathbb{E}\,|e_s|^{2q}\, \Delta t + c\Delta t^{2q}.
$$

Since finally

$$
c \lceil T/\Delta t\rceil^{q-1} \sum_{s=0}^{\ell-1} \mathbb{E}|R_{s+1}|^{2q} \leq c\Delta t^{2q},
$$

by inserting these three estimates in (20) we end up with

$$
\mathbb{E}\left[\sup_{k=0,\ldots,\ell} |e_k|^{2q}\right] \leq c \sum_{s=0}^{\ell-1} \mathbb{E}\,|e_s|^{2q}\, \Delta t + c\Delta t^{2q}
$$

and Gronwall's Lemma completes the proof.

$\square$

The above result and Lemma 2.5 now give convergence (without a rate) in all L^q-norms:

Corollary 2.8. *Under the assumptions of Theorem 2.7 we have*

$$\lim_{\Delta t \to 0} \mathbb{E} \left[\sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |x(k\Delta t) - X_k|^q \right] = 0 \qquad (23)$$

for all $q \geq 1$.

Remark 2.9. *In [2], independently of the research in this paper, a similar result to Theorem 2.7 is established for the case $D = (\alpha, \infty)$ and drift functions $f : D \to \mathbb{R}$, which are twice continuously differentiable and satisfy a monotone condition (which is equivalent to our one-sided Lipschitz condition). Using a continuous extension of BEM Alfonsi obtains the error bound (15) under the assumption*

$$\mathbb{E} \left(\int_0^T |f'(x(t))|^2 \, dt \right)^{p/2} + \mathbb{E} \left(\int_0^T \left| (f'f)(x(t)) + \frac{\sigma^2}{2} f''(x(t)) \right| dt \right)^p < \infty$$

for $p \geq 1$. This result is then applied to the CIR and CEV process, i.e. Propositions 3.1 and 3.3 are obtained.

Note that due to our bound on the inverse moments on the LBE, see the subsection below, we are also able to cover SDEs like the the Heston 3/2-volatility and the Ait-Sahlia model. Moreover, since we work under the assumption $D = (\alpha, \beta)$ we can also treat the Wright-Fisher SDE and similar equations.

2.4 Boundedness of inverse moments of BEM

If the drift coefficient has an even more specific structure, see the assumption below, we can also control the inverse moments of BEM. For the Heston-3/2 volatility and also the Ait-Sahalia model this will be helpful later on.

Assumption 2.10. *Let $\alpha \geq 0$ and assume that the drift coefficient $f : (\alpha, \beta) \to (\alpha, \beta)$ has the structure*

$$f(x) = \frac{c_1}{x^{m_1}} + h(x), \qquad x \in (\alpha, \beta)$$

where

$$|h(x)| \leq c_2 \cdot (1 + |x|^{m_2}) \qquad x \in (\alpha, \beta)$$

for some $c_1, c_2 > 0$ and $m_1, m_2 > 0$.

Under the above assumption BEM can be written as

$$X_{k+1} = X_k + (c_1 X_{k+1}^{-m_1} + h(X_{k+1}))\Delta t + \sigma \Delta w_{k+1}$$

and we have

$$\frac{1}{X_{k+1}^{m_1}} = \frac{1}{c_1 \Delta t} \left(X_{k+1} - X_k - h(X_{k+1})\Delta t - \sigma \Delta w_{k+1} \right). \qquad (24)$$

Proceeding as in the proof of Theorem 2.7 we also have

$$\frac{1}{x(t_{k+1})^{m_1}} = \frac{1}{c_1 \Delta t} \left(x(t_{k+1}) - x(t_k) - h(x(t_{k+1}))\Delta t - \sigma \Delta w_{k+1} - R_{k+1} \right), \quad (25)$$

with R_{k+1} given by (17). This can be used to derive the following result:

Lemma 2.11. *Let $T > 0$ and $p \geq 2$. Moreover, let the assumptions of Theorem 2.7 and let also Assumption 2.10 hold. Then there exists constants $C_p^{(1)}, C_p^{(2)} > 0$ such that we have*

$$\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} \mathbb{E}|X_k|^{-m_1 p} < C_p^{(1)}\left(1 + \sup_{t\in[0,T]} \mathbb{E}|x(t)|^{-m_1 p}\right).$$

and

$$\mathbb{E}\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |X_k|^{-m_1 p} < C_p^{(2)}\left(1 + \mathbb{E}\sup_{t\in[0,T]} |x(t)|^{-m_1 p}\right).$$

Proof. We only prove the second assertion, the proof of the first assertion is similar. Using (24), (25), (15), (21) and (22) we have

$$\mathbb{E}\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |x(t_k)^{-m_1} - X_k^{-m_1}|^p \leq c\left(1 + \mathbb{E}\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |h(x(t_k)) - h(X_k)|^p\right).$$

Since

$$|h(x) - h(y)| \leq |h(x)| + |h(y)| \leq c_2(1 + |x|^{m_2} + |y|^{m_2}),$$

Lemma 2.5 and the triangle inequality finish the proof. $\square$

In the next Lemma we establish a general a-priori estimate for uniform inverse moments of SDE (6) with the drift structure imposed by Assumption 2.10.

Lemma 2.12. *Let $p \geq 2$. Let the assumptions of Lemma 2.5 hold and in addition let the drift of SDE (6) satisfy Assumption 2.10. Then there exists a constant $C_p > 0$ such that we have*

$$\mathbb{E}\sup_{t\in[0,T]} |x(t)|^{-p} \leq C_p\left(1 + \sup_{t\in[0,T]} \mathbb{E}|x(t)|^{-(p+2)}\right).$$

Proof. Let $\{\alpha_n\}_{n\in\mathbb{N}}$ and $\{\beta_n\}_{n\in\mathbb{N}}$ be such that $\alpha_n \searrow \alpha$ and $\beta_n \nearrow \beta$, when $n \to \infty$. Let us choose n_0 such that $\alpha_{n_0} \leq x(0) \leq \beta_{n_0}$. Then for $n \geq n_0$, we define the stopping time $\tau_n = \{t > 0 : x(t) \notin (\alpha_n, \beta_n)\}$. By Itô's lemma we have

$$|x(t \wedge \tau_n)|^{-p} = |x(0)|^{-p} + \int_0^t \left(g(x(s)) - p|x(s)|^{-(p+1)}h(x(s))\right)\mathbf{1}_{[0,\tau_n]}(s)ds$$

$$- p\sigma \int_0^t |x(s)|^{-(p+1)}\mathbf{1}_{[0,\tau_n]}(s)dw(s),$$

where $g(x) = -c_1 p|x|^{-(p+1+m_1)} + \frac{\sigma^2 p(p+1)}{2}|x|^{-(p+2)}$. Observe that for $m_1 > 1$, $\lim_{x\searrow 0} g(x) = -\infty$ and $\lim_{x\to\infty} g(x) = 0$, hence in that case there exists a $c > 0$ such that

$$\sup_{x>0} g(x) \leq c.$$

If $m_1 \in (0,1]$ then there exists a $c > 0$ such that

$$|g(x)| \leq c(1 + |x|^{-(p+2)}).$$

By Assumption 2.10 and Burkholder-Davis-Gundy's inequality we have

$$\mathbb{E}\left[\sup_{t\in[0,T]}|x(t\wedge\tau_n)|^{-p}\right] \leq |x(0)|^{-p} + \int_0^T c\left(1 + \mathbb{E}|x(s)|^{-(p+2)} + \mathbb{E}|x(s)|^{-(p+1)+m_2}\right)ds$$

$$+ c\,\mathbb{E}\left(\int_0^T |x(s)|^{-2(p+1)}\mathbf{1}_{[0,\tau_n]}(s)ds\right)^{1/2}$$

$$\leq |x(0)|^{-p} + \int_0^T c\left(1 + \mathbb{E}|x(s)|^{-(p+2)}\right)ds$$

$$+ c\,\mathbb{E}\left(\sup_{t\in[0,T]}|x(t\wedge\tau_n)|^{-p}\int_0^T |x(s)|^{-(p+2)}ds\right)^{1/2}.$$

Applying Young's inequality to the last summand of the above inequality now yields

$$\mathbb{E}\left[\sup_{t\in[0,T]}|x(t\wedge\tau_n)|^{-p}\right] \leq c\left(1 + \sup_{t\in[0,T]}\mathbb{E}|x(t)|^{-(p+2)}\right).$$

and Fatou's Lemma completes the proof.

$\square$

3 Examples

In this section, we will apply our main result to several examples. To simplify the presentation, we will denote the numerical method $Y_k = F^{-1}(X_k)$, $k = 0, 1, \ldots$, where X_k, $k = 0, 1, \ldots$ is given by (5), as *Lamperti-backward Euler* (LBE) approximation of SDE (4). Moreover, we will say that this method is *p-strongly convergent with order one*, if

$$\mathbb{E}\sup_{k=0,\ldots,\lceil T/\Delta t\rceil}|y(t_k) - Y_k|^p \leq C_p \cdot \Delta t^p.$$

Finally, constants whose particular value is not important will be again denoted by c.

3.1 CIR process

Recall that the Cox-Ingersoll-Ross process is given by the SDE

$$dy(t) = \kappa(\theta - y(t))dt + \sigma\sqrt{y(t)}dw(t), \quad t \geq 0, \qquad y(0) > 0. \qquad (26)$$

If $2\kappa\theta \geq \sigma^2$, then we have $D = (0, \infty)$ and Assumption (2.1) holds for $(\alpha, \beta) = (0, \infty)$. Moreover, recall that the transformed SDE using $F(y) = \sqrt{y}$ reads as

$$dx(t) = f(x(t))dt + \frac{1}{2}\sigma dw(t), \quad t \geq 0, \qquad x(0) = \sqrt{y(0)} \qquad (27)$$

with

$$f(x) = \frac{1}{2}\kappa\left(\left(\theta - \frac{\sigma^2}{4\kappa}\right)x^{-1} - x\right), \qquad x > 0$$

13

and the BEM scheme is given by

$$X_{k+1} = X_k + f(X_{k+1})\Delta t + \frac{1}{2}\sigma\Delta w_{k+1}, \qquad k = 0, 1, \dots \qquad (28)$$

with $X_0 = x(0)$. Straightforward calculations give

$$(x - y)(f(x) - f(y)) \leq -\frac{1}{2}\kappa(x - y)^2, \qquad x, y > 0,$$

so Assumption 2.2 holds with $K = -\kappa/2$. Observe also that

$$f'(x) = -\frac{1}{2}\kappa(\theta_v x^{-2} + 1),$$

and

$$(f'f)(x) + \frac{\sigma^2}{2}f''(x) = -\frac{\kappa^2}{4}(\theta_v^2 x^{-3} - x) + \frac{1}{2}\kappa\theta_v x^{-3}\sigma^2.$$

So, for Assumption 2.6 to hold we need

$$\sup_{0 \leq t \leq T} \mathbb{E}[x(t)^{-3p}] = \sup_{0 \leq t \leq T} \mathbb{E}[y(t)^{-\frac{3}{2}p}].$$

Since

$$\sup_{0 \leq t \leq T} \mathbb{E}[y(t)^q] < \infty \quad \text{for} \quad q > -\frac{2k\theta}{\sigma^2}, \qquad (29)$$

see e.g. [6], Assumption 2.6 and as a consequence Theorem 2.7 hold if $p < \frac{4}{3}\frac{k\theta}{\sigma^2}$.

In order to approximate the original CIR process observe that

$$(x(t_k)^2 - X_{t_k}^2) = (x(t_k) + X_{t_k})(x(t_k) - X_{t_k}).$$

Let $\varepsilon > 0$ such that $p(1 + \varepsilon) < \frac{4}{3}\frac{k\theta}{\sigma^2}$. Then Hölder's inequality gives

$$\mathbb{E}\left[\sup_{0 \leq k \leq \lceil T/\Delta t \rceil} \left|x(t_k)^2 - X_{t_k}^2\right|^p\right]$$

$$\leq \left(\mathbb{E}\left[\sup_{0 \leq k \leq \lceil T/\Delta t \rceil} |x(t_k) + X_{t_k}|^{p\frac{1+\varepsilon}{\varepsilon}}\right]\right)^{p\frac{\varepsilon}{1+\varepsilon}} \left(\mathbb{E}\left[\sup_{0 \leq k \leq \lceil T/\Delta t \rceil} |x(t_k) - X_{t_k}|^{p(1+\varepsilon)}\right]\right)^{\frac{1}{1+\varepsilon}}.$$

Using Lemma 2.5 we obtain:

Proposition 3.1. *Let $T > 0$ and $2 \leq p < \frac{4}{3}\frac{\kappa\theta}{\sigma^2}$. Then, the LBE approximation of the CIR process is p-strongly convergent with order one.*

3.2 Numerical Experiment

Note that the unique solution to (28) is given by

$$X_{k+1} = \frac{1}{2 + \kappa\Delta t}\left(X_k + \frac{1}{2}\sigma\Delta w_{k+1} + \sqrt{\left(X_k + \frac{1}{2}\sigma\Delta w_{k+1}\right)^2 + \kappa\theta_v\Delta t}\right)$$

14

with $\theta_v = \theta - \frac{\sigma^2}{4\kappa}$. Hence implicit schemes not necessarily increase the computational complexity with comparison to classical explicit procedures.

In our numerical experiment, we focus on the L^2-error at the endpoint $T = 1$, so we let

$$e_{\Delta t} = \mathbb{E} \left| x(T) - X_T \right|^2 .$$

For our numerical experiment we set $\theta = 0.125$, $\kappa = 2$, and $\sigma = 0.5$. This gives $\frac{2\kappa\theta}{\sigma^2} = 2$ and corresponds to the critical parameters for which Dereich et al. [6] established strong convergence of order one half for linearly interpolated BEM (28) with respect to the uniform L^2-error criteria. Although, theoretical results obtained in this paper impose slightly more restrictive conditions for parameters than those in Dereich et al., performed numerical experiment suggests that for practical simulations condition $1 < p < \frac{2\kappa\theta}{\sigma^2}$ suffices for uniform-L^p convergence with order one. Although an explicit solution to (27) is unknown, Theorem 2.7 guarantees that BEM strongly converges to the true solution. Therefore, it is reasonable to take BEM with a very small time step, we choose $\Delta t = 2^{-15}$, as a reference solution. We then compare it to BEM evaluated with $(2^4 \Delta t, 2^5 \Delta t, 2^6 \Delta t, 2^7 \Delta t)$ in order to estimate the rate of the L^2-convergence, where we estimate $e_{\Delta t}$ by a Monte-Carlo procedure, i.e.

$$e_{\Delta t} \approx \frac{1}{10^3} \sum_{i=1}^{10^3} \left| x^{(i)}(T) - X_T^{(i)} \right|^2 .$$

Here $x^{(i)}(T), X_T^{(i)}$ are iid copies of $x(T), X_T$. We plot $e_{\Delta t}$ against Δt on a log-log scale, i.e. if we assume that a power law relation $e_{\Delta t} = C\Delta t^q$ holds for some constant C and q, then we have $\log e_{\Delta t} = \log C + q \log \Delta t$. For our simulation, a least squares fit for $\log C$ and q yields the value 1.9332 for q with a least square residual of 0.016. Hence, our results are consistent with strong order of convergence equal to one.
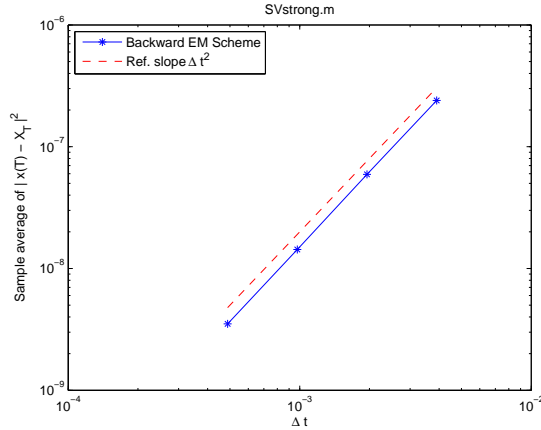

Figure 1: Strong error plot for backward Euler scheme applied to CIR process.

15

3.3 Heston 3/2-volatility

In [10] the inverse of a CIR process is used as a stochastic volatility process, which gives the so-called Heston-3/2-volatility

$$dy(t) = c_1 y(t)(c_2 - y(t))\, dt + c_3 y(t)^{3/2}\, dw(t), \qquad t \geq 0, \qquad y(0) > 0 \quad (30)$$

where $c_1, c_2, c_3 > 0$. Using $F(y) = y^{-1/2}$ leads to

$$dx(t) = \left(\left(\frac{c_1}{2} + \frac{3c_3^2}{8} \right) x(t)^{-1} - \frac{c_1 c_2}{2} x(t) \right) dt - \frac{c_3}{2}\, dw(t), \qquad (31)$$

which coincides with (27) if we use a reflected Brownian motion, i.e. $-w$, which is still a Brownian motion, and

$$\sigma = c_3, \qquad \theta = \frac{1}{c_2} + \frac{c_3^2}{c_1 c_2}, \qquad \kappa = c_1 c_2.$$

Hence we have the relation

$$\frac{\kappa \theta}{\sigma^2} = 1 + \frac{c_1}{c_3^2},$$

so Theorem 2.7 holds here for $p < \frac{4}{3}(1 + \frac{c_1}{c_3^2})$. Note that the Heston-3/2-volatility is one of the SDEs which does not have finite moments of any order. As the inverse of the CIR process it has finite q-moments up to order $q < 2 + \frac{2c_1}{c_3^2}$.

Now, for transforming back we have to control the inverse moments of the BEM scheme for CIR. Here Lemma 2.11 and 2.12 give

$$\mathbb{E} \sup_{k=0,\dots,\lceil T/\Delta t \rceil} |X_k|^{-p} \leq C_p \left(1 + \sup_{t \in [0,T]} \mathbb{E}|x(t)|^{-(p+2)} \right)$$

for $p < \frac{4}{3}(1 + \frac{c_1}{c_3^2})$. From the analysis of the CIR process we have that

$$\sup_{t \in [0,T]} \mathbb{E}|x(t)|^{-(p+2)} < \infty \quad \text{for} \quad p < 4\left(\frac{1}{2} + \frac{c_1}{c_3^2} \right).$$

To establish the convergence result for the LBM for the Heston-3/2 volatility note that

$$\left| \frac{1}{X_k^2} - \frac{1}{x(t_k)^2} \right| = \frac{X_k + x(t_k)}{X_k^2 x(t_k)^2} |X_k - x(t_k)| \leq c(|X_k|^{-3} + |x(t_k)|^{-3})|X_k - x(t_k)|.$$

Finally using Hölder's inequality gives

$$\mathbb{E}\left[\sup_{k=0,\dots,\lceil T/\Delta t \rceil} \left| \frac{1}{X_k^2} - \frac{1}{x(t_k)^2} \right|^p \right]$$

$$\leq c \left(\left(\mathbb{E} \sup_{k=0,\dots,\lceil T/\Delta t \rceil} |X_k|^{-4p} \right)^{3/4} + \left(\mathbb{E} \sup_{k=0,\dots,\lceil T/\Delta t \rceil} |x(t_k)|^{-4} \right)^{3/4} \right)$$

$$\times \left(\mathbb{E} \sup_{k=0,\dots,\lceil T/\Delta t \rceil} |X_k - x(t_k)|^{4p} \right)^{1/4}.$$

Hence we obtain:

16

Proposition 3.2. *If* $1 \le p < \frac{1}{3} + \frac{1}{3}\frac{c_1}{c_3^2}$, *then the LBE approximation of the Heston-3/2 process is p-strongly convergent with order one.*

3.4 CEV process

Another popular SDE in finance is the mean reverting constant elasticity of variance process ([5]) given by

$$dy(t) = \kappa(\theta - y(t))dt + \sigma y(t)^\alpha dw(t) \tag{32}$$

where $0.5 < \alpha < 1$, $\kappa, \theta, \sigma > 0$. By the Feller test we have $D = (0, \infty)$. Applying Itô's formula to the function $F(y(t)) = y(t)^{1-\alpha}$ we obtain that Assumptions 2.1 holds with $(\alpha, \beta) = (0, \infty)$ and

$$dx(t) = f(x(t))dt + (1 - \alpha)\sigma dw(t) \tag{33}$$

with

$$f(x) = (1 - \alpha)\left(\kappa\theta x^{-\frac{\alpha}{1-\alpha}} - \kappa x - \frac{\alpha\sigma^2}{2}x^{-1}\right), \qquad x > 0.$$

Again we need to check the remaining assumptions. Since $\alpha > 0.5$, and consequently $\frac{1}{1-\alpha} > 2$, we have for

$$f'(x) = -\alpha\kappa\theta x^{-\frac{1}{1-\alpha}} - (1 - \alpha)\kappa + (1 - \alpha)\frac{\alpha\sigma^2}{2}x^{-2}, \qquad x > 0$$

that

$$\lim_{x \to 0} f'(x) = -\infty, \qquad \lim_{x \to \infty} f'(x) = -(1 - \alpha)\kappa$$

and hence there exists a $c > 0$ such that

$$\sup_{x>0} f'(x) \le c.$$

Now the mean value theorem implies

$$(x - y)(f(x) - f(y)) \le c\,|x - y|^2, \qquad x, y > 0,$$

i.e. the drift coefficient is one-sided Lipschitz. We have moreover that

$$f''(x) = \frac{\alpha}{1 - \alpha}\kappa\theta x^{-\frac{2-\alpha}{1-\alpha}} - (1 - \alpha)\alpha\sigma^2 x^{-3}.$$

However, from [4] it is known that

$$\sup_{0 \le t \le T} \mathbb{E}\,|y(t)|^p < \infty \qquad \text{for any} \quad p \in \mathbb{R},\ T > 0, \tag{34}$$

and therefore also Assumption 2.6 holds and Theorem 2.7 can be applied for any $p \ge 1$. For the back transformation note that the mean value theorem yields

$$|(x^{\frac{1}{1-\alpha}} - y^{\frac{1}{1-\alpha}})| \le \frac{1}{1 - \alpha}(x^{\frac{\alpha}{1-\alpha}} + y^{\frac{\alpha}{1-\alpha}})|x - y|,$$

Using Lemma 2.5 we have:

Proposition 3.3. *Let $p \ge 1$. The LBE approximation of the CEV process is p-strongly convergent with order one.*

3.5 Wright-Fisher Diffusion

The Wright-Fisher SDE that originated from mathematical biology,[7] and recently is also gaining popularity in mathematical finance [14, 25] reads as

$$dy(t) = (a - by(t))dt + \gamma\sqrt{|y(t)(1 - y(t))|}\, dw(t), \qquad t \geq 0, \qquad y_0 \in (0, 1)$$

with $a, b, \gamma > 0$. If

$$\frac{2a}{\gamma^2} \geq 1 \qquad \text{and} \qquad \frac{2(b - a)}{\gamma^2} \geq 1, \tag{35}$$

then this SDE has a unique strong solution with

$$\mathbb{P}(y(t) \in (0, 1),\ t \geq 0) = 1,$$

see [28]. Using

$$F(y) = 2\arcsin(\sqrt{y}), \qquad y \in (0, 1)$$

we obtain

$$dx(t) = f(x(t))dt + \gamma\, dw(t), \qquad t \geq 0, \qquad x_0 = 2\arcsin(\sqrt{y_0})$$

with

$$f(x) = \left(a - \frac{\gamma^2}{4}\right)\cot\left(\frac{x}{2}\right) - \left(b - a - \frac{\gamma^2}{4}\right)\tan\left(\frac{x}{2}\right), \qquad x \in (0, \pi).$$

Since

$$f'(x) = -\frac{1}{2}\left(a - \frac{\gamma^2}{4}\right)\left(1 + \cot^2\left(\frac{x}{2}\right)\right) - \frac{1}{2}\left(b - a - \frac{\gamma^2}{4}\right)\left(1 + \tan^2\left(\frac{x}{2}\right)\right)$$

the mean value theorem implies that Assumption 2.2 is satisfied with $K = 0$. Now note that

$$f''(x) = \frac{1}{2}\left(a - \frac{\gamma^2}{4}\right)\cot\left(\frac{x}{2}\right)\left(1 + \cot^2\left(\frac{x}{2}\right)\right)$$
$$- \frac{1}{2}\left(b - a - \frac{\gamma^2}{4}\right)\tan\left(\frac{x}{2}\right)\left(1 + \tan^2\left(\frac{x}{2}\right)\right)$$

and

$$(f'f)(x(t)) + \frac{\gamma^2}{2}f''(x(t)) = (f'f)(2\arcsin(y(t))) + \frac{\gamma^2}{2}f''(2\arcsin(y(t))).$$

Since

$$\cot(\arcsin(y(t))) = \sqrt{\frac{1 - y(t)}{y(t)}}, \qquad \tan(\arcsin(y(t))) = \sqrt{\frac{y(t)}{1 - y(t)}}$$

and $y(t) \in (0, 1)$ we obtain

$$|f'(x(t))| \leq c\left(1 + |y(t)|^{-1} + |1 - y(t)|^{-1}\right)$$

and

$$\left|(f'f)(x(t)) + \frac{\gamma^2}{2}f''(x(t))\right| \leq c\left(1 + |y(t)|^{-3/2} + |1 - y(t)|^{-3/2}\right)$$

18

for some constant $c > 0$, depending only on $a, b, \gamma > 0$. Using Theorem 4.1 in [15] and establishing uniform convergence of the given series expressions in t using asymptotic bounds on the Jacobi polynomials we have that

$$\sup_{t \in [0,T]} \mathbb{E}|y(t)|^{q_1} < \infty$$

if $q_1 > -\frac{2a}{\gamma^2}$ and

$$\sup_{t \in [0,T]} \mathbb{E}|1 - y(t)|^{q_2} < \infty$$

if $q_2 > -\frac{2(b-a)}{\gamma^2}$. So Assumption 2.6 is satisfied if

$$\frac{4}{3\gamma^2} \min\{a, b - a\} > p.$$

Now Theorem 2.7 gives

$$\mathbb{E}\left[\sup_{k=0,\ldots,\lceil T/\Delta t \rceil} |x(t_k) - X_k|^p \right] \leq C_p \cdot \Delta t^p \qquad \text{for} \qquad p < \frac{4}{3\gamma^2} \min\{a, b - a\}$$

and transforming back yields

$$\mathbb{E}\left[\sup_{k=0,\ldots,\lceil T/\Delta t \rceil} \left| y(t_k) - \sin^2\left(\frac{X_k}{2}\right) \right|^p \right] \leq C_p \cdot \Delta t^p.$$

under the same assumption on the parameters since the sin-function is bounded.

Proposition 3.4. *Let $2 \leq p < \frac{4}{3\gamma^2} \min\{a, b - a\}$. Then the LBE approximation of the Wright-Fisher process is p-strongly convergent with order one.*

3.6 Ait-Sahalia model

Higham et al. analysed in [29] a backward Euler scheme for the Ait-Sahalia interest rate model

$$dy(t) = (\alpha_{-1} y(t)^{-1} - \alpha_0 + \alpha_1 y(t) - \alpha_2 y(t)^r)dt + \sigma y(t)^\rho dw(t), \qquad (36)$$

where $\alpha_{-1}, \alpha_0, \alpha, \alpha_1, \alpha_2, \sigma$ are positive constants and $\rho, r > 1$. In [29] it was established that

$$\mathbb{P}(y(t) \in (0, \infty), \ t \geq 0) = 1.$$

Under the assumption $r + 1 > 2\rho$ Higham et al. proved uniform L^p-convergence for any $p \geq 2$ the backward Euler method directly applied to (36). However, their results did not reveal a rate of convergence. Here, using the Lamperti transformation approach we construct a scheme that strongly converges with rate one. We focus on the critical case with $r = 2$ and $\rho = 1.5$ which was not covered in [29]. Using $F(y) = y^{-1/2}$ we obtain

$$dx(t) = f(x)dt - \frac{1}{2}\sigma dw(t).$$

with

$$f(x) = \left(\frac{1}{2}\alpha_2 + \frac{3}{8}\sigma^2\right)x(t)^{-1} - \frac{1}{2}\alpha_1 x(t) + \frac{1}{2}\alpha_0 x(t)^3 - \frac{1}{2}\alpha_{-1}x(t)^5, \quad x > 0.$$

We have for

$$f'(x) = -\left(\frac{1}{2}\alpha_2 + \frac{3}{8}\sigma^2\right)x(t)^{-2} - \frac{1}{2}\alpha_1 + \frac{3}{2}\alpha_0 x(t)^2 - \frac{5}{2}\alpha_{-1}x(t)^4,$$

that

$$\lim_{x \to 0} f'(x) = -\infty, \qquad \lim_{x \to \infty} f'(x) = -\infty$$

and hence there exists a $c > 0$ such that

$$\sup_{x > 0} f'(x) < c.$$

Now the mean value theorem implies

$$(x - y)(f(x) - f(y)) \le c\,|x - y|^2\,, \qquad x, y > 0,$$

i.e. the drift coefficient is one-sided Lipschitz. We have moreover that

$$f''(x) = \left(\alpha_2 + \frac{3}{4}\sigma^2\right)x(t)^{-3} + 3\alpha_0 x(t) - 10\,\alpha_{-1}x(t)^3,$$

and

$$\left|(f'f)(x(t)) + \frac{\sigma^2}{8}f''(x(t))\right| \le c\left(1 + |x(t)|^{-3} + |x(t)|^{20}\right).$$

Straightforward computations also give that

$$f(x) > g(x), \qquad x \in (0, 2x(0)]$$

where

$$g(x) := \left(\frac{1}{2}\alpha_2 + \frac{3}{8}\sigma^2\right)x^{-1} - \frac{1}{2}\alpha_2\beta x$$

with

$$\beta = \frac{\alpha_1}{\alpha_2} + 16\frac{\alpha_{-1}}{\alpha_2}x(0)^4.$$

Now a comparison result for SDEs, see e.g. Proposition V.2.18 and Exercise V.2.19 in [22], yields that almost surely

$$x(t) \ge x^{(1)}(t), \qquad t \in [0, \tau^{(1)}),$$

where

$$dx_n^{(1)}(t) = g(x^{(1)}(t))dt - \frac{1}{2}\sigma dw(t), \quad t \ge 0, \qquad x^{(1)}(0) = x(0)$$

and

$$\tau^{(1)} = \inf\{t \in [0, T] : x(t) > 2x(0)\}.$$

Let us define a sequence of stopping times

$$\tau^{(2i)} = \inf\{t \in [\tau^{(2i-1)}, T] : x(t) \le 2x(0)\}, \quad i = 1, 2, \ldots,$$

$$\tau^{(2i+1)} = \inf\{t \in [\tau^{(2i)}, T] : x(t) > 2x(0)\}, \quad i = 1, 2, \dots$$

and an associated sequence of SDEs

$$dx^{(i)}(t) = g(x^{(i)}(t))dt - \frac{1}{2}\sigma dw(t), \quad t \geq \tau^{(2i)}, \qquad x^{(i)}(\tau^{(2i)}) = x(\tau^{(2i)}).$$

Using the comparison result for SDEs again, we have that almost surely

$$x(t) \geq x^{(i)}(t), \qquad t \in [\tau^{(2i)}, \tau^{(2i+1)}).$$

Therefore, using the CIR process as a lower bound, Assumption 2.6 and consequently Theorem 2.7 holds for $p < \frac{4}{3}\left(1 + \frac{\alpha_2}{\sigma^2}\right)$. Proceeding as for the 3/2-model we have:

Proposition 3.5. *Let $1 \leq p < \frac{1}{3} + \frac{1}{3}\frac{\alpha_2}{\sigma^2}$. The LBE approximation of the Ait-Sahalia process with $r = 2$ and $\rho = 1.5$ is p-strongly convergent with order one.*

In the case $r + 1 > 2\rho$ we know from [29] that

$$\sup_{t \in [0,T]} \mathbb{E}\,|x(t)|^{-p} < \infty \qquad \text{for all} \qquad p \geq 1.$$

Moreover, the drift coefficient of the transformed SDE behaves at zero like the one of a transformed CEV process. A by now standard analysis gives:

Proposition 3.6. *Let $p \geq 1$. The LBE approximation of the Ait-Sahalia process with $r + 1 > 2\rho$ is p-strongly convergent with order one.*

4 A Milstein-type scheme for CIR

In this section we establish a connection between the Lamperti-backward Euler and a drift-implicit Milstein scheme for the CIR process. We will show that the order of convergence of the LBE carries over to a drift-implicit Milstein scheme, which has been proposed in [20] and [13]. While strong convergence was shown in [23], sharp convergence rates have not been established so far.

Recall that BEM for the transformed CIR process reads as

$$X_{k+1} = X_k + f(X_{k+1})\Delta t + \frac{1}{2}\sigma \Delta w_{k+1}, \qquad k = 0, 1, \dots$$

with

$$f(x) = \frac{1}{2}\kappa\left(\left(\theta - \frac{\sigma^2}{4\kappa}\right)x^{-1} - x\right), \qquad x > 0.$$

Squaring yields the LBE, i.e.

$$X_{k+1}^2 = X_k^2 + \kappa(\theta - X_{k+1}^2)\Delta t + \sigma X_k \Delta w_{k+1}$$
$$+ \frac{\sigma^2}{4}\left((\Delta w_{k+1})^2 - \Delta t\right) - \left(f(X_{k+1})\right)^2 \Delta t^2.$$

On the other hand the drift-implicit Milstein scheme for CIR is given by

$$Z_{k+1} = Z_k + \kappa(\theta - Z_{k+1})\Delta t + \sigma\sqrt{Z_k}\Delta w_{k+1} + \frac{\sigma^2}{4}\left((\Delta w_{k+1})^2 - \Delta t\right), \quad (37)$$

hence both schemes coincide up to a term of order Δt^2. The numerical flows of the LBE and Milstein scheme are given by

$$\phi_E(x, k) = \frac{1}{1 + \kappa \Delta t} \left(x + \kappa \theta \Delta t + \sigma \sqrt{x} \Delta w_{k+1} + \frac{\sigma^2}{4} \left((\Delta w_{k+1})^2 - \Delta t \right) \right)$$
$$- \frac{1}{1 + \kappa \Delta t} \left(f(\phi_E(x, k))^2 \Delta t^2 \right.$$

and

$$\phi_M(x, k) = \frac{1}{1 + \kappa \Delta t} \left(x + \kappa \theta \Delta t + \sigma \sqrt{x} \Delta w_{k+1} + \frac{\sigma^2}{4} \left((\Delta w_{k+1})^2 - \Delta t \right) \right).$$

It is clear then that

$$\phi_M(x, k) \geq \phi_E(x, k)$$

for all $x > 0$, $k = 0, 1, \ldots$. From [1] we know on the other hand

$$\phi_E(x, k) \geq \phi_E(y, k) \quad \text{for} \quad x \geq y.$$

Hence we conclude

$$Z_k \geq X_k^2, \quad k = 0, 1, \ldots, \tag{38}$$

so the drift-implicit Milstein scheme dominates the Lamperti-Euler method and thus preserves positivity.

To establish the order of L^1-convergence for the drift-implicit Milstein scheme it is enough to control the difference between the Lamperti-Euler method and (37).

Lemma 4.1. Let $\frac{\kappa \theta}{\sigma^2} > 3/2$. Then there exists a constant $c > 0$ such that

$$\sup_{k=0,\ldots,\lceil T/\Delta t \rceil} \mathbb{E} \left| Z_k - X_k^2 \right| \leq c \Delta t. \tag{39}$$

Proof. Let $e_k = Z_k - X_k^2$ and note that $e_k \geq 0$ by (38). We have

$$e_{k+1} = e_k - \kappa e_{k+1} \Delta t + \sigma (\sqrt{Z_k} - X_k) \Delta w_{k+1} + (f(X_{k+1}))^2 \Delta t^2.$$

Exploiting the independence of X_k, Z_k and Δw_{k+1} it follows

$$\mathbb{E} e_{k+1} = \frac{1}{1 + \kappa \Delta t} \left(\mathbb{E} e_k + \mathbb{E}(f(X_{k+1}))^2 \Delta t^2 \right)$$

and consequently

$$\sup_{k=0,\ldots,\lceil T/\Delta t \rceil} \mathbb{E} e_k \leq \Delta t \sum_{k=0}^{\lceil T/\Delta t \rceil - 1} \mathbb{E}(f(X_{k+1}))^2 \Delta t.$$

Due to our assumptions Lemma 2.11 gives that

$$\sup_{k=1,\ldots,\lceil T/\Delta t \rceil} \mathbb{E} X_k^{-2} \leq c,$$

which together with

$$|f(x)| \leq c \cdot \left(1 + |x| + |x|^{-1} \right)$$

and Lemma 2.5 shows the assertion.

$\square$

Using this result we have:

Proposition 4.2. *(i) Let $\frac{\kappa\theta}{\sigma^2} > 3/2$. Then, there exists a constant $C > 0$ such that*

$$\sup_{k=0,\ldots,\lceil T/\Delta t\rceil} \mathbb{E}|y(k\Delta t) - Z_k| \leq C \cdot \Delta t. \tag{40}$$

(ii) Let $\frac{\kappa\theta}{\sigma^2} > 3/2$. Then, there exists a constant $C > 0$ such that

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |y(k\Delta t) - Z_k|^2 \leq C \cdot \Delta t. \tag{41}$$

Proof. (i) This follows from the triangle inequality and Proposition 3.1.

(ii) Using (40) the second assertion can be shown along the lines of the proof of Proposition 5.3 in [23], where strong convergence of the drift-implicit Milstein scheme (without a convergence rate) was shown. Proceeding as in the proof of Proposition 5.3 in [23] we have

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |y(k\Delta t) - Z_k|^2 \leq c \sum_{\ell=0}^{\lceil T/\Delta t\rceil - 1} (1 + \kappa\Delta t)^{2\ell} \mathbb{E}\,|y(\ell\Delta t) - Z_\ell|\,\Delta t$$

$$+ c\,\mathbb{E} \sup_{k=1,\ldots,\lceil T/\Delta t\rceil} \left| \sum_{\ell=0}^{k-1} \frac{1}{(1 + \kappa\Delta t)^{k-\ell}} \rho_{\ell+1} \right|^2$$

with

$$\rho_{k+1} = -\kappa \int_{k\Delta t}^{(k+1)\Delta t} (y(s) - y((k+1)\Delta t))\,ds + \sigma \int_{k\Delta t}^{(k+1)\Delta t} (\sqrt{y(s)} - \sqrt{y(k\Delta t)})\,dw(s).$$

So, (40) gives

$$\mathbb{E} \sup_{k=0,\ldots,\lceil T/\Delta t\rceil} |y(k\Delta t) - Z_k|^2 \leq c\Delta t + c\,\mathbb{E} \sup_{k=1,\ldots,\lceil T/\Delta t\rceil} \left| \sum_{\ell=0}^{k-1} \frac{1}{(1 + \kappa\Delta t)^{k-\ell}} \rho_{\ell+1} \right|^2.$$

For the second term straightforward computations yield that

$$\mathbb{E} \sup_{k=1,\ldots,\lceil T/\Delta t\rceil} \left| \sum_{\ell=0}^{k-1} \frac{1}{(1 + \kappa\Delta t)^{k-\ell}} \rho_{\ell+1} \right|^2 \leq c\Delta t,$$

which completes the proof of the proposition. $\qquad\square$

Using the (suboptimal) second estimate of the above Proposition and Lemma 3.5 of [6] we also obtain a sharp error estimate for the piecewise linear interpolation of the drift-implicit Milstein scheme, i.e.

$$\overline{Z}_t = \frac{t_{k+1} - t}{\Delta} Z_k + \frac{t - t_k}{\Delta} Z_{k+1}, \qquad t \in [t_k, t_{k+1}],$$

in a combined L^2-$\|\cdot\|_\infty$ norm.

Proposition 4.3. *Let $\frac{\kappa\theta}{\sigma^2} > 3/2$. Then, there exists a constant $C > 0$ such that*

$$\left(\mathbb{E} \max_{t \in [0,T]} |y(t) - \overline{Z}_t|^2 \right)^{1/2} \leq C \cdot \sqrt{|\log(\Delta t)|} \cdot \sqrt{\Delta t},$$

for all $\Delta \in (0, 1/2]$.

The above relation between a drift-implicit Milstein scheme applied to the original SDE

$$dy(t) = a(y(t))dt + b(y(t))dw(t)$$

and the BEM applied to the transformed SDE

$$dx(t) = f(x(t))dt + \lambda dw(t)$$

with

$$f(x) = \lambda \left(\frac{a(F^{-1}(x))}{b(F^{-1}(x))} - \frac{1}{2}b'(F^{-1}(x)) \right)$$

and

$$F(x) = \lambda \int^x \frac{1}{b(y)} dy$$

is in fact a particular case of a more general relation. Expanding LBE yields

$$F^{-1}(X_{k+1}) = F^{-1}(X_k) + \frac{1}{\lambda}b(F^{-1}(X_k))(X_{k+1} - X_k)$$
$$+ \frac{1}{2\lambda^2}b'(F^{-1}(X_k))b(F^{-1}(X_k))(X_{k+1} - X_k)^2 + \dots,$$

since

$$\frac{d}{dx}F^{-1}(x) = \frac{1}{\lambda}b(F^{-1}(x)), \qquad \frac{d^2}{(dx)^2}F^{-1}(x) = \frac{1}{\lambda^2}b'(F^{-1}(x))b(F^{-1}(x)).$$

Setting $Y_k = F^{-1}(X_k)$ and using (5) we have

$$Y_{k+1} = Y_k + \left(a(Y_{k+1}) - \frac{1}{2}b'b(Y_{k+1}) \right) \Delta t + b(Y_k)\Delta w_{k+1} + \frac{1}{2}b'b(Y_k)|\Delta w_{k+1}|^2$$
$$+ R_k$$

with

$$R_k = (b(Y_k) - b(Y_{k+1}))f(X_{k+1})\Delta t + \frac{1}{\lambda}(b'b)(Y_k)\Delta w_{k+1}f(X_{k+1})\Delta t$$
$$+ \frac{1}{2\lambda^2}(b'b)(Y_k)f^2(X_{k+1})\Delta t^2 + \dots$$

So dropping R_k and the other higher order terms, we end up with

$$Z_{k+1} = Z_k + \left(a(Z_{k+1}) - \frac{1}{2}b'b(Z_{k+1}) \right) \Delta t + b(Z_k)\Delta w_{k+1} + \frac{1}{2}b'b(Z_k)|\Delta w_{k+1}|^2.$$

In the case of $(\alpha, \beta) = (0, \infty)$ conditions for the well-definedness, stability and strong convergence of this scheme are given in [13]. However, the convergence rate analysis for the CIR process, where we can exploit (among other things) that $F^{-1}(x) = x^2$ and also the domination property (38), seems not to carry over to the general case.

5 Conclusion and Discussion

In this paper we presented a Lamperti-Euler scheme for scalar SDEs which take values in a domain $D = (l, r)$ and have non-Lipschitz drift or diffusion coefficients. Our strategy is to first use the Lamperti transformation $x(t) = F(y(t))$ (provided that the diffusion coefficient of the original SDE is strictly positive on D) and then to approximate the transformed process $x(t)$, $0 \leq t \leq T$, with the backward Euler scheme. Transforming back with the inverse Lamperti transformation gives an approximation scheme for the original SDE. We also pointed out a relation of this scheme to a drift-implicit Milstein scheme.

The advantages of this Lamperti-Euler method are

- that it preserves the domain of the original SDE

- and an available framework which allows to establish strong convergence order one for this scheme.

In particular, we use this framework to obtain such strong convergence results for several SDEs with non-Lipschitz coefficients from both mathematical finance and bio-mathematics.

Whether the implicitness of this scheme (which e.g. for the CEV process requires solving a non-linear equation) can be avoided by using a tamed Euler scheme (as in [17] for the case $D = \mathbb{R}$) remains an open question. Open questions are also, whether Assumption 2.2 and 2.6 can be formulated in terms of conditions on the original coefficients of the SDE, and whether the convergence rate for the Lamperti-Euler also carries over in general (and not only for the CIR process) to a drift-implicit Milstein scheme. In particular, the last point leads to the general question that given a certain numerical approximation (for an SDE with non-Lipschitz coefficients) which perturbations do not change its convergence properties and also its qualitative properties? We will pursue all these topics in our future research.

References

[1] A. Alfonsi. On the discretization schemes for the CIR (and Bessel squared) processes. *Monte Carlo Methods and Applications*, 11(4):355–384, 2005.

[2] A. Alfonsi. Strong convergence of some drift implicit Euler scheme. Application to the CIR process. *Working Paper*, 2012.

[3] L.B.G. Andersen. Simple and Efficient Simulation of the Heston Stochastic Volatility Model. *Journal of Computational Finance*, 11(3):1–42, 2008.

[4] A. Berkaoui, M. Bossy, and A. Diop. Euler scheme for SDEs with non-Lipschitz diffusion coefficient: strong convergence. *ESAIM: Probability and Statistics*, 12:1–11, 2007.

[5] J. Cox. Notes on option pricing I: Constant elasticity of variance diffusions. *Unpublished manuscript (Stanford University, Stanford, CAL)*, 1975.

[6] S. Dereich, A. Neuenkirch, and L. Szpruch. An euler-type method for the strong approximation of the cox–ingersoll–ross process. *Proceedings of the Royal Society A Engineering Science*, 468(2140):1105–1115, 2012.

[7] S. Either and T. Kurtz. *Markov Processes: Characterization and Convergences.* John Wiley Sons, New York, 1986.

[8] M. Giles. Improved multilevel Monte Carlo convergence using the Milstein scheme. *Monte Carlo and Quasi-Monte Carlo Methods 2006, Proceedings*, pages 343–358, 2006.

[9] M.B. Giles. Multilevel monte carlo path simulation. *Operations Research-Baltimore*, 56(3):607–617, 2008.

[10] S.L. Heston. A simple new formula for options with stochastic volatility. *Course notes of Washington University in St. Louis, Missouri*, 1997.

[11] D.J. Higham and X. Mao. Convergence of Monte Carlo simulations involving the mean-reverting square root process. *Journal of Computational Finance*, 8(3):35, 2005.

[12] D.J. Higham, X. Mao, and A.M. Stuart. Strong convergence of Euler-type methods for nonlinear stochastic differential equations. *SIAM Journal on Numerical Analysis*, 40(3):1041–1063, 2003.

[13] D.J. Higham, X. Mao, and L. Szpruch. Convergence, non-negativity and stability of a new milstein scheme with applications to finance. *Arxiv preprint arXiv:1204.1647*, 2012.

[14] S. Howison and D. Schwarz. Risk-neutral pricing of financial instruments in emission markets. *to appear in SIFIN*, 2012.

[15] T.R. Hurd and A. Kuznetsov. Explicit formulas for Laplace transforms of stochastic integrals. *Markov Processes and Related Fields*, 14(2):277–290, 2008.

[16] Jentzen A. Hutzenthaler, M. and P.E. Kloeden. Divergence of the multilevel monte carlo euler method for nonlinear stochastic differential equations. *Arxiv preprint arXiv:1105.0226*, 2011.

[17] Jentzen A. Hutzenthaler, M. and P.E. Kloeden. Strong convergence of an explicit numerical method for sdes with non–globally lipschitz continuous coefficients. *Annals of Applied Probability*, 22(4):1611 – 1641, 2012.

[18] S.M. Iacus. *Simulation and inference for stochastic differential equations: with R examples.* Springer Verlag, 2008.

[19] A. Jentzen, P.E. Kloeden, and A. Neuenkirch. Pathwise approximation of stochastic differential equations on domains: higher order convergence rates without global Lipschitz coefficients. *Numerische Mathematik*, 112(1):41–64, 2009.

[20] C. Kahl, M. Günther, and T. Rosberg. Structure preserving stochastic integration schemes in interest rate derivative modeling. *Applied Numerical Mathematics*, 58(3):284–295, 2008.

[21] C. Kahl and P. Jackel. Fast strong approximation Monte Carlo schemes for stochastic volatility models. *Quantitative Finance*, 6(6):513–536, 2006.

[22] I. Karatzas and S.E. Shreve. *Brownian Motion and Stochastic Calculus*. Springer, 1991.

[23] P.E. Kloeden and A. Neuenkirch. Convergence of Numerical Methods for Stochastic Differential Equations in Mathematical Finance. *Working Paper*, 2012.

[24] P.E. Kloeden and E. Platen. *Numerical Solution of Stochastic Differential Equations*. Springer, 1992.

[25] K.S. Larsen and M. Sorensen. Diffusion models for exchange rates in a target zone. *Mathematical Finance*, 17(2):285–306, 2007.

[26] R. Lord, R. Koekkoek, and D.J.C. Van Dijk. A comparison of biased simulation schemes for stochastic volatility models. *Quantitative Finance*, 10(2):177–194, 2009.

[27] X. Mao and L. Szpruch. Strong convergence rates for backward euler–maruyama method for non-linear dissipative-type stochastic differential equations with super-linear diffusion coefficients. *Stochastics*, to appear.

[28] Karlin S. and Taylor H.M. *A Second Course in Stochastic Processes*. Academic Press, 1981.

[29] L. Szpruch, X. Mao, D.J. Higham, and J. Pan. Strongly nonlinear ait-sahalia-type interest rate model and its numerical approximation. *BIT Numerical Mathematics*, 51(2):405–425, 2010.

[30] E. Zeidler. *Nonlinear Functional Analysis and its Applications*. Springer Verlag, 1985.

[31] J. Zhu. *Modular Pricing of Options: An Application of Fourier Analysis*. Springer Verlag, 2009.